PROCÉDÉ DE CREATION IN VITRO DE SEQUENCES POLYNUCLEOTIDIQUES RECOMBINEES PAR LIGATION ORIENTEE

La présente invention se rapporte à une méthode d'obtention *in vitro* de séquences polynucléotidiques recombinées par ligation orientée. L'invention vise tout particulièrement à générer puis sélectionner des séquences polynucléotidiques susceptibles de présenter une ou plusieurs propriétés avantageuses par rapport aux propriétés correspondantes de séquences de référence et donc capables de conférer un phénotype amélioré et/ou de produire des protéines améliorées.

On entend par séquence de référence une séquence ayant des propriétés proches de celles recherchées.

On entend par *in vitro* tout événement, réaction ou procédé qui n'a pas lieu dans un organisme vivant.

On entend par ligation un procédé qui permet de créer une liaison phosphodiester entre deux fragments d'acides nucléiques.

On entend par ligation orientée tout procédé de ligation qui permet d'assembler dans un ordre déterminé des molécules d'acides nucléiques, notamment par hybridation desdites molécules d'acides nucléiques sur au moins une matrice nucléotidique.

On entend par séquence polynucléotidique toute molécule d'acide nucléique simple ou double brin.

Différentes techniques ont été développées pour favoriser la recombinaison *in vitro* entre différentes séquences polynucléotidiques, parmi celles-ci on peut citer plus particulièrement le DNA-Shuffling et le StEP toutes deux fondées sur l'utilisation de la PCR.

Le DNA-Shuffling comporte deux étapes, la fragmentation aléatoire par la DNAse I de séquences polynucléotidiques, et une amplification par PCR dans laquelle les fragments précédemment générés servent d'amorces. A chaque étape d'hybridation, le changement de

matrice provoque des recombinaisons au niveau des régions ayant des séquences homologues. Le StEP consiste à mélanger différentes séquences polynucléotidiques contenant diverses mutations en présence d'une paire d'amorces. Ce mélange est soumis à une réaction de type PCR dans laquelle les étapes d'hybridation et de polymérisation sont regroupées en une seule étape de très courte durée. Ces conditions permettent l'hybridation des amorces mais diminuent la vitesse de polymérisation, de façon à ce que les fragments qui sont partiellement synthétisés s'hybrident aléatoirement sur les séquences polynucléotidiques portant les différentes mutations, permettant ainsi la recombinaison. Dans chacune de ces deux méthodes, l'étape de polymérisation est indispensable au processus de recombinaison. Ainsi, en fonction des polymérases choisies, cette étape de polymérisation peut être génératrice de mutations supplémentaires non souhaitées. En outre, à partir d'un certain nombre de cycles, le DNA-Shuffling et le StEP reposent sur le principe de l'hybridation d'une "méga-amorce" sur une matrice, ce qui entraîne probablement des difficultés de mise en oeuvre pour des séquences polynucléotidiques dont la taille est supérieure à 1,5 Kpb. Enfin, ces deux techniques ne permettent pas de contrôler le taux de recombinaisons, puisque celles-ci se font aléatoirement au cours des étapes successives de polymérisation.

La présente invention vise précisément à palier les inconvénients ci-dessus en offrant une méthode simple de préparation d'au moins une séquence polynucléotidique recombinée, en utilisant un procédé de ligation orientée de fragments obtenus à partir d'une banque de séquences polynucléotidiques.

On entend par banque de séquences polynucléotidiques un ensemble de séquences polynucléotidiques contenant au moins deux séquences polynucléotidiques hétérologues.

Dans une forme de mise en œuvre de l'invention, ce but est atteint grâce à un procédé comprenant les étapes suivantes :

a) la fragmentation d'une banque de séquences polynucléotidiques ,

b) la dénaturation des fragments ainsi obtenus,

c) l'hybridation de fragments obtenus à l'étapes (b) avec une ou plusieurs matrice(s) d'assemblage,

d) la ligation orientée desdits fragments pour obtenir au moins une séquence polynucléotidique recombinée,

Lorsque la matrice d'assembalge est double brin, elle est préalablement dénaturée avant l'étape (c) comme par exemple au cours de l'étape (b).

Le procédé de l'invention permet de recombiner de manière aléatoire différents fragments au cours des étapes (b), (c) et (d) au sein d'une séquence polynucléotidique. Ce procédé reproduit donc *in vitro* les phénomènes de recombinaison qui peuvent se produire *in vivo* en les favorisant. Le procédé de l'invention est donc tout particulièrement intéressant pour recombiner des séquences polynucléotidiques entre elles afin de générer de nouvelles séquences polynucléotidiques

Ces séquences polynucléotidiques recombinées sont susceptibles de présenter des propriétés avantageuses par rapport aux propriétés correspondantes de séquences de référence et donc capables de conférer un phénotype amélioré et/ou de produire des protéines améliorées.

Ce but est atteint grâce à un procédé comprenant les étapes suivantes :

a) la fragmentation d'une banque de séquences polynucléotidiques ,

b) la dénaturation des fragments,

c) l'hybridation de fragments obtenus à l'étape (b) avec une ou plusieurs matrice(s) d'assemblage,

d) la ligation orientée desdits fragments pour obtenir au moins une séquence polynucléotidique recombinée,

e) la sélection des séquences polynucléotidiques recombinées présentant des propriétés avantageuses par rapport aux propriétés correspondantes à une ou plusieurs séquences de référence.

La ou les matrice(s) d'assemblage peuvent-être simple ou double brin. Dans le cas ou l'une de ces matrices est double brin, elle est préalablement dénaturée pour être ajoutée sous forme simple brin à l'étape (c) comme par exemple au cours de l'étape (b).

Le procédé de l'invention peut comprendre à l'issue de l'étape (e) la répétition des étapes (a), (b), (c) et (d). Dans ce cas, la banque de séquences polynucléotidiques contient au moins une séquence polynucléotidique recombinée qui a été sélectionnée en (e).

Le procédé de l'invention peut aussi comprendre à l'issue de l'étape (d) et avant l'étape (e), la répétition des étapes (b), (c) et (d), ou des étapes (a), (b), (c) et (d).

Ce dernier mode de réalisation est particulièrement utile dans le cas où à l'issue de l'étape (d) tous les fragments ne sont pas ligués. Dans ce cas, le procédé de l'invention comprend en outre, à la fin de l'étape (d) et avant l'étape (e), un ou plusieurs cycles des réactions suivantes:

- dénaturation des fragments ligués et non ligués issus de l'étape (d), éventuellement en présence d'une ou plusieurs matrice(s) d'assemblage,
- hybridation desdits fragments avec une ou plusieurs matrice(s) d'assemblage si celle(s)-ci n'est (ne sont) pas présente(s) lors de la dénaturation,
- ligation desdits fragments.

Ces réactions de dénaturation, hybridation et ligation, sont équivalentes aux étapes (b), (c) et (d), mais réalisées non pas avec les fragments de l'étape (a)

mais avec les fragments ligués et non ligués issus de l'étape (d).

Selon une forme de mise en œuvre paticulière du procédé, les séquences polynucléotidiques de la banque sont simple brin. L'usage de fragments d'ADN simple brin est particulièrement adapté à la recombinaison de familles de gènes pour lesquelles une matrice simple brin donnée ou un mélange de matrices simple brin différentes seront hybridées à des fragments simple brin issus d'une banque de gènes homologues. Puisqu'il n'y a pas de séquences strictement complémentaires entre elles parmi la population de fragments, l'hybridation ne sera pas biaisée vers les séquences sauvages entre les fragments ou entre fragments et matrices. La température d'hybridation peut alors être ajustée en fonction du degré d'homologie entre les séquences, générant ainsi des molécules recombinées avec le degré de recombinaison le plus élevé possible. Des banques de molécules recombinées sont ainsi créées, avec une meilleure valeur en terme de diversité, augmentant considérablement les chances de trouver le bon mutant a l'issue du minimum de cycles de recombinaison.

Pour obtenir des molécules simple brin d'ADN, un phagemide de type Bluescript ou un vecteur de la famille des phages filamenteux comme M13mp18 peuvent être utilisés. Une autre façon de procéder consiste à générer des molécules double brin par PCR en utilisant une amorce phosphorylée en 5' et l'autre non phosphorylée. La digestion par l'exonucléase du phage lambda détruira les brins d'ADN phosphorylés en 5', laissant les brins non phosphorylés intacts. Une autre méthode de génération de molécules simple brin consiste à faire une amplification par PCR assymétrique à partir d'une matrice d'ADN méthylée.

La digestion par *Dpn* I détruira les brins méthylés, laissant intacts les produits de l'amplification qui pourront alors être purifiés après dénaturation.

Le procédé de l'invention peut comprendre en outre, une ou plusieurs des étapes suivantes :

- la séparation des séquences polynucléotidiques recombinées de la ou des matrice(s) d'assemblage avant l'étape (e),
- l'amplification des séquences polynucléotidiques recombinées avant l'étape (e).
- le clonage de séquences polynucléotidiques recombinées éventuellement après séparation des brins recombinés de la ou des matrices

Dans une forme de mise en œuvre avantageuse de la méthode, les extrémités des fragments générés à l'étape (a) sont telles qu'il peut y avoir hybridation adjacente de ces extrémités au moins à une matrice d'assemblage à l'étape (c) et ligation de ces fragments les uns avec les autres à l'étape (d). Les séquences polynucléotidiques de la banque sur laquelle le procédé de l'invention est effectué doivent être telles que les fragments obtenus au cours du procédé présentent des extrémités telles que décrites ci-dessus. Ces fragments peuvent être notamment obtenus au cours de l'étape (a), ou au cours de l'étape (d) par ligation de fragments.

Une forme de mise en oeuvre avantageuse du procédé de l'invention consiste à réaliser simultanément les étapes (c) et (d) selon une réaction dite de RLR pour l'expression anglaise de "Recombining Ligation Reaction".

Outre les avantages indiqués précédemment, le procédé de l'invention est remarquable en ce qu'il favorise et accélère la recombinaison aléatoire *in vitro* de séquences polynucléotidiques, ces séquences polynucléotidiques pouvant être des gènes. On entend par gène un fragment ou une séquence d'ADN associée à une fonction biologique. Un gène peut être obtenu de différentes façons, dont la synthèse chimique, la synthèse par polymérisation ou par extraction dudit gène à partir d'une source d'acides nucléiques.

La recombinaison *in vitro* des séquences polynucléotidiques de la banque initiale par le procédé de l'invention permet donc d'obtenir une nouvelle banque contenant des séquences ayant acquis une ou plusieurs caractéristiques des séquences de la banque précédente. Le procédé de l'invention constitue donc une technique d'évolution *in vitro*.

Le procédé de l'invention constitue une alternative à la PCR recombinatoire telle que mise en oeuvre dans les techniques de DNA shuffling ou de StEP, puisqu'il ne nécessite pas d'étape de polymérisation *in vitro* pour aboutir à la recombinaison. Au contraire, l'étape clef du procédé de l'invention est l'étape (d) de ligation sur une matrice d'assemblage (ou ligation orientée), ce qui assure un très haut degré de fidélité au cours des événements de recombinaison.

Le procédé de l'invention est remarquable en ce qu'il permet d'augmenter considérablement l'efficacité du réassemblage des fragments à liguer en utilisant la ligation orientée. En effet, dans le cas d'une séquence

découpée en n fragments, il existe de très nombreuses possibilités de réassociation des fragments en utilisant un procédé de ligation classique (sans utilisation d'une matrice de réassemblage qui oriente la ligation), parmi lesquelles une seule forme est intéressante. Dans le cas du procédé de l'invention, la ligation est orientée par la matrice d'assemblage, ce qui permet d'obtenir directement la seule forme intéressante.

La fragmentation de ces séquences polynucléotidiques à l'étape (a) peut se faire soit de manière contrôlée, soit de manière aléatoire.

Dans le cas d'une fragmentation réalisée de manière contrôlée, la fragmentation permet de maîtriser avec précision le degré de recombinaison voulu et la position des points de recombinaison. Selon une forme de réalisation préférée du procédé de l'invention, l'étape (a) consiste à soumettre les séquences polynucléotidiques de la banque à une hydrolyse par l'action d'une ou plusieurs enzymes de restriction. Ainsi, dans une forme de mise en œuvre particulière du procédé de l'invention, le degré de recombinaison et la position des points de recombinaison des séquences polynucléotidiques recombinées sont déterminés par la fragmentation de l'étape (a).

Ainsi, plus le nombre de fragments générés par séquence est grand, plus le nombre de fragments nécessaires pour recomposer une séquence est élevé, ce qui entraîne un taux de recombinaison élevé. En outre, la nature et la position des extrémités des fragments générés dans ce mode de réalisation du procédé de l'invention peuvent être connues et contrôlées, ce qui permet :

- de contrôler avec précision les zones où la recombinaison a lieu, ou

- d'induire la recombinaison entre des séquences polynucléotidiques, par exemple des gènes, si les extrémités des fragments sont créées dans des zones d'homologie entre ces séquences, ou des zones d'homologies entre ces séquences et la ou les matrices d'assemblage.

Dans le cas d'une fragmentation aléatoire, tout moyen enzymatique ou mécanique connu de l'homme de métier capable de couper aléatoirement l'ADN pourra être utilisé, comme par exemple une digestion par la Dnase I ou une sonication.

Le procédé de l'invention permettant d'augmenter considérablement l'efficacité de réassemblage des fragments à liguer, il peut donc être appliqué à l'orientation de la ligation multi-moléculaire à bouts francs. Dans cette application, on utilise comme matrice d'assemblage aux étapes (b) ou (c) des oligonucléotides simple ou double brin juste complémentaires de l'extrémité 3' d'un fragment et 5' du fragment adjacent, ce qui permet l'hybridation adjacente de ces deux extrémités sur la même matrice après l'étape de dénaturation. Une fois hybridées, les extrémités des fragments peuvent être liguées entre-elles de façon à orienter le sens de ligation des fragments à bout francs. La même approche peut être envisagée pour l'orientation de la ligation de fragments à bouts cohésifs.

Une mise en œuvre toute préférée du procédé de l'invention consiste à ajouter des enzymes capables, à l'étape (c) et/ou à l'étape (d), de reconnaître et de

dégrader et/ou couper de manière spécifique les extrémités non hybridées de fragments, lorsque celles-ci recouvrent d'autres fragments hybridés sur la même matrice. Un exemple préféré de ce type d'enzyme est l'enzyme Flap endonucléase .

Une mise en œuvre particulière du procédé de l'invention consiste donc à utiliser des enzymes du type Flap endonucléases lorsque les fragments générés à l'étape (a) peuvent se recouvrir lors de l'hybridation sur la matrice d'assemblage à l'étape (c).

Ainsi, lors de l'hybridation de fragments d'ADN sur une matrice, ces enzymes ont pour propriété de reconnaître et de couper de manière spécifique les extrémités non hybridées de ces fragments, lorsque celles-ci recouvrent d'autres fragments hybridés sur la même matrice.

Lorsque les fragments utilisés au cours du procédé de l'invention sont double brin, une forme particulière de l'invention consiste à utiliser des enzymes de type exonucléase spécifiques du simple brin. Ces enzymes auront pour propriété de reconnaître et de dégrader de manière spécifique les extrémités simple brin non hybridées de ces fragments, lorsque celles-ci recouvrent d'autres fragments hybridés sur la même matrice.

Au cours de l'étape (c) d'hybridation, l'utilisation de ce type d'enzymes (notamment Flap, ou exonucléase spécifique du simple brin) permet donc d'augmenter le nombre d'extrémités adjacentes pouvant être liguées à l'étape (d), ce qui est particulièrement

important dans le cas de fragments obtenus par coupure aléatoire, car ces fragments présentent des zones de recouvrement les uns avec les autres lorsqu'ils s'hybrident sur la matrice d'assemblage.

Dans une mise en œuvre particulière du procédé de l'invention utilisant une ligase active à haute température et préférentiellement thermostable à l'étape (d), les enzymes capables de reconnaître et/ou de couper de manière spécifique les extrêmités non hybridées des fragments, ajoutées à l'étape (c) et/ou à l'étape (d) auront les mêmes propriétés de thermorésistance et d'activité à haute température que ladite ligase.

La banque de séquences polynucléotidiques sur laquelle est effectuée le procédé de l'invention peut être générée par toute méthode connue de l'homme du métier, par exemple à partir d'un gène sauvage par étapes de mutagénèse dirigée successives, par PCR "error prone " (2), par mutagénèse aléatoire chimique, par mutagénèse aléatoire *in vivo*, ou en combinant des gènes de familles proches ou distinctes au sein de la même espèce ou d'espèces différentes de façon à disposer d'une variété de séquences polynucléotidiques dans ladite banque.

Parmi ces techniques, l'invention envisage plus particulièrement, un procédé dans lequel la banque de séquences polynucléotidiques est obtenue par une réaction de polymérisation en chaîne réalisée dans des conditions qui permettent de créer des mutations ponctuelles aléatoires.

La banque initiale de séquences polynucléotidiques peut être constituée de séquences

synthétiques qui seront fragmentées à l'étape (a) ou qui peuvent constituer les fragments de l'étape (a).

Selon une forme de réalisation préférée du procédé de l'invention, l'étape (a) consiste à soumettre les séquences polynucléotidiques de la banque à une hydrolyse par l'action d'une ou plusieurs enzymes de restriction.

Pour accroître le degré de recombinaison généré par le procédé de l'invention, il suffit d'augmenter le nombre de fragments de restriction en utilisant des enzymes de restriction ayant un grand nombre de sites de coupure sur les séquences polynucléotidiques de la banque, ou en combinant plusieurs enzymes de restriction. Dans le cas de l'utilisation d'une ligase thermostable et thermoactive, la taille du plus petit fragment ainsi généré sera avantageusement supérieure ou égale à 40 b ou 40 pb, afin de conserver une température d'hybridation compatible avec celle de l'étape (d) de ligation qui est généralement de l'ordre de 65 °C.

L'étape (a) peut encore être réalisée en générant une banque de fragments par traitement aléatoire enzymatique ou mécanique. En particulier, l'étape (a) peut consister en un traitement aléatoire avec la DNAse I d'une banque de séquences polynucléotidiques . Dans le cas où l'on utilise à l'étape (a) une fragmentation enzymatique ou mécanique aléatoire, cette forme de mise en oeuvre du procédé de l'invention a la particularité de permettre l'utilisation des fragments générés par ce traitement comme matrices les uns pour les autres, pour l'hybridation au cours de l'étape (c) ou de la réaction de RLR des étapes (c) et (d) simultanées.

L'étape (b) peut être réalisée en combinant au moins deux banques de fragments distinctes générées séparément à l'étape (a) à partir de la même banque initiale par des traitements différents, comme par exemple avec des enzymes de restriction différentes. Dans le cas de la mise en oeuvre de telles banques, on utilise les fragments obtenus à l'étape (a) comme matrices les uns pour les autres, pour l'hybridation au cours de l'étape (c) ou de la réaction de RLR des étapes (c) et (d) simultanées.

Les fragments de l'étape (a) du procédé de l'invention peuvent également être générés par des réactions d'amplification (telle que la PCR) menées sur les séquences polynucléotidiques de la banque. Deux solutions sont notamment envisageables. Dans un premier cas, les oligonucléotides amorces peuvent être conçus de manière à générer des fragments dont les extrémités sont adjacentes tout au long de la séquence d'assemblage. Dans un deuxième cas, les oligonucléotides amorces sont conçus de façon à générer des fragments ayant des séquences communes, ces fragments pouvant servir de matrice d'assemblage les uns pour les autres à l'étape (b) ou à l'étape (c).

L'efficacité de recombinaison du procédé de l'invention est fonction du nombre de fragments générés par séquence polynucléotidique à l'étape (a). En conséquence, le procédé de l'invention utilisera des séquences polynucléotidiques ayant été fragmentés en n fragments, n étant avantageusement supérieur ou égal à trois.

La matrice d'assemblage à l'étape (b) ou (c) est par exemple une séquence polynucléotidique issue de la banque initiale ou une séquence consensus de ladite banque, simple ou double brin. Dans le cas où la ou les matrices d'assemblage sont incorporées directement à l'étape (c) de l'invention, cette matrice doit être sous forme simple brin.

Selon une variante du procédé de l'invention, les matrices d'assemblage des étapes (b) ou (c) sont constituées d'oligonucléotides simples ou doubles brins.

Selon une forme particulière de mise en oeuvre du procédé de l'invention, des oligonucléotides, simple ou double brin, de longueur variable, sont ajoutés à l'étape (b) ou (c) en plus de la matrice. Ces oligonucléotides sont conçus pour pouvoir se substituer à une partie des fragments à l'étape (c), en effet, leur séquence est telle que :

- s'ils sont parfaitement homologues avec la séquence du fragment qu'ils remplacent, ils favorisent certaines combinaisons, ou

- s'ils sont partiellement hétérologues avec la séquence du fragment qu'ils remplacent, ils introduisent une ou des mutations dirigées supplémentaires.

On entend par séquences hétérologues, deux séquences dont la composition en bases diffère d'au moins une base.

Avant l'étape (e) du procédé de l'invention, il est possible de séparer les séquences polynucléotidiques recombinées de la matrice d'assemblage grâce à un marqueur présent sur la matrice d'assemblage ou sur les séquences polynucléotidiques recombinées. Il est en effet possible de

marquer chaque brin de la matrice selon des techniques connues de l'homme du métier. Par exemple, le marqueur de la matrice d'assemblage peut être un haptène et l'on sépare les séquences polynucléotidiques recombinées de la matrice d'assemblage par des techniques connues de l'homme du métier, comme par exemple un anticorps anti-haptène fixé sur un support ou une réaction biotine-streptavidine, si l'haptène est un marqueur biotine.

D'autres techniques peuvent être employées pour séparer les séquences polynucléotidiques recombinées de la matrice d'assemblage. La matrice d'assemblage peut aussi être préparée spécifiquement de façon à faciliter son élimination à la fin du procédé de l'invention. Elle peut ainsi être synthétisée par amplification PCR utilisant des dATP méthylés, ce qui permet sa dégradation par l'endonucléase de restriction *Dpn* I. Dans ce cas, les séquences polynucléotidiques recombinées ne doivent pas contenir de dATP méthylés. La matrice peut aussi avoir été préparée par amplification PCR en utilisant des dUTP, ce qui permet sa dégradation par traitement avec une uracyl-DNA-glycosylase. A l'inverse, il est possible de protéger les séquences polynucléotidiques recombinées en les amplifiant par PCR sélective avec des oligonucléotides portant des groupements phosphorothioates en 5'. Un traitement avec une exonucléase permet alors de dégrader spécifiquement la matrice d'assemblage.

Le procédé de l'invention peut comprendre avant le clonage éventuel de l'étape (e), une étape d'amplification des séquences polynucléotidiques recombinées. Toute technique d'amplification est acceptable notamment une amplification par PCR. Une des plus simple

consiste à réaliser une PCR qui permet d'amplifier spécifiquement les séquences polynucléotidiques recombinées grâce à des amorces qui ne peuvent s'hybrider que sur les extrémités des séquences recombinées. Les produits PCR sont ensuite clonés pour être caractérisés et les séquences polynucléotidiques présentant des propriétés avantageuses par rapport aux propriétés correspondantes de séquences de référence sont sélectionnées.

L'invention a pour objet de générer des séquences polynucléotidiques susceptibles de présenter des propriétés avantageuses par rapport aux propriétés correspondantes de séquences de référence. Les séquences polynucléotides recombinées obtenues à l'étape (d) et éventuellement clonées sont criblées par tout moyen approprié pour sélectionner les séquences polynucléotidiques recombinées ou les clones présentant des propriétés avantageuses par rapport aux propriétés correspondantes de séquences de référence. On entend, par exemple, par propriétés avantageuses la thermostabilité d'une enzyme ou sa qualité à pouvoir fonctionner dans des conditions de pH ou de température ou de concentration saline plus adaptées à un procédé enzymatique, que les protéines témoins habituellement utilisées pour ledit procédé. A titre d'exemple d'un tel procédé, on peut citer un procédé industriel de désencollage des fibres textiles ou de blanchiment des pâtes à papier ou de la production d'arômes dans l'industrie laitière, les procédés de biocatalyse pour la synthèse par voie enzymatique de nouvelles molécules thérapeutiques, etc…

Selon un mode de réalisation avantageux du procédé de l'invention, la banque de séquence

polynucléotidique peut donc être le résultat d'un crible ayant permis de sélectionner par tout moyen approprié les séquences polynucléotidiques présentant des propriétés avantageuses par rapport à des séquences témoins. Les séquences ainsi sélectionnées constituent une banque restreinte.

Mais, il est aussi possible de partir d'une banque non restreinte afin de conserver la représentativité des propriétés contenues dans cette banque.

Les séquences codant pour la ou les protéines présentant une ou des propriétés avantageuses par rapport aux protéines de référence sont alors sélectionnées, par des criblages *in vivo* ou *in vitro*, et peuvent servir à constituer une nouvelle banque pour une éventuelle réitération du procédé de l'invention. Une mise en œuvre avantageuse du procédé de l'invention consiste donc à utiliser comme banque plusieurs séquences polynucléotidiques sélectionnées après une première mise en œuvre du procédé de l'invention, éventuellement mélangées avec d'autres séquences polynucléotidiques. Parmi les techniques de criblage qui peuvent être appliquées à chacun des clones de l'étape (e), les techniques de criblage par expression *in vitro* utilisant notamment la transcription *in vitro* des séquences polynucléotidiques recombinées puis la traduction *in vitro* des mRNAs obtenus, présentent l'avantage de s'affranchir des problèmes de physiologie cellulaire, et de tous les inconvénients liés au clonage d'expression *in vivo*. En outre, ce type de criblage est facilement automatisable, ce qui permet de cribler un nombre élevé de séquences polynucléotidiques recombinées.

L'invention concerne aussi une séquence polynucléotidique recombinée obtenue par un procédé selon l'invention, ainsi qu'un vecteur contenant une telle séquence polynucléotidique recombinée, un hôte cellulaire transformé par une séquence polynucléotidique recombinée ou un vecteur de l'invention, ainsi qu'une protéine codée par cette séquence polynucléotidique recombinée. L'invention comprend également les banques correspondantes de séquences polynucléotidiques recombinées, de vecteurs, d'hôtes cellulaires ou de protéines.

Exemple :

Cet exemple a pour objectif de générer des séquences polynucléotidiques recombinées du gène de résistance à la kanamycine en utilisant la ligation orientée de fragments simples brin

Dans un premier temps, le gène de résistance à la kanamycine (1 Kb) de pACYC184 est cloné dans le polylinker de M13mp18 de telle façon que le phagemide simple brin contient le brin non codant du gène.

En parallèle, ce gène est amplifié par une PCR mutagène (error prone PCR) avec deux amorces qui sont complémentaires de la séquence du vecteur M13mp18 de par et d'autre de la séquence du gène. L'amorce pour le brin non codant est phosphorylée alors que l'amorce pour le brin codant ne l'est pas. Le produit de la PCR mutagène est digéré par l'exonucléase de lambda, ce qui génère une banque de brins codants de mutants du gène de la résistance à la kanamycine.

Cette banque de molécules simple brin est digérée par un mélange d'enzymes de restriction, à savoir Hae III, Hinf I et Taq I.

Cette banque de fragments simple brin ainsi obtenue est alors hybridée au phagemide simple brin et liguaturée avec une ligase thermostable. Cette étape est répétée plusieurs fois, jusqu'à ce que les fragments de petite taille ne puissent plus être observés lors d'un dépôt sur un gel d'agarose alors que la bande correspondant au simple brin du gène complet de la résistance à la kanamycine devienne un composant majeur du " smear " de molécules simple brin visibles sur le gel.

La bande correspondant à la taille du gène est alors découpée du gel et purifiée. Elle est ensuite hybridée avec deux oligonucléotides (40 mer) complémentaires des séquences de M13mp18 de chaque côté du gène, et ce duplex partiel est digéré par Eco RI et Sph I puis ligaturée dans un vecteur M13 mp18 digéré par les mêmes enzymes.

Les cellules transformées avec le produit de ligation sont criblées pour une résistance accrue à la kanamycine.

Le clonage des molécules simple brin recombinées peut éventuellement être réalisé par une PCR avec deux amorces du gène de taille complète et clonage du produit double brin de cette amplification. Pour éviter les mutations indésirables, cette amplification sera réalisée avec la polymérase de type Pfu et par un nombre limité de cycles.

Les plasmides des clones significativement plus résistants à la kanamycine que la souche initiale sont purifiés et utilisés comme matrices pour une PCR à la polymérase Pfu, dans des conditions de haute fidélité, avec le couple d'amorces phospshorylé/non phosphorylé tel que défini précédemment. Ceci génère la seconde génération de fragments simple brin, après un traitement à l'exonucléase de lambda et la fragmentation avec les enzymes de restriction. Les enzymes utilisées à cette étape peuvent être composées d'un mélange différent (Bst NI, Taq I et Mnl I).

Les étapes de recombinaison et de sélection sont répétées plusieurs fois, jusqu'à ce qu'un accroissement substantiel de la résistance à la kanamycine soit obtenu.